Article

Stationary Schrödinger Equation and Darwin Term from Maximal Entropy Random Walk

Manfried Faber

Atominstitut, Technische Universität Wien, Stadionallee 2, 1020 Wien, Austria; faber@kph.tuwien.ac.at

Abstract: We describe particles in a potential by a special diffusion process, the maximal entropy random walk (MERW) on a lattice. Since MERW originates in a variational problem, it shares the linear algebra of Hilbert spaces with quantum mechanics. The Born rule appears from measurements between equilibrium states in the past and the same equilibrium states in the future. Introducing potentials by the observation that time, in a gravitational field running in different heights with a different speed, MERW respects the rule that all trajectories of the same duration are counted with equal probability. In this way, MERW allows us to derive the Schrödinger equation for a particle in a potential and the Darwin term of the nonrelativistic expansion of the Dirac equation. Finally, we discuss why quantum mechanics cannot be simply a result of MERW, but, due to the many analogies, MERW may pave the way for further understanding.

Keywords: Schrödinger equation; Darwin term; diffusion; interpretation of quantum mechanics

1. Introduction

The phenomena at atomic and subatomic scales are very successfully described by quantum mechanics and the Schrödinger equation [1]. Despite existing now for almost 100 years, their interpretation is still a topic of intensive discussion [2–5]. We know how to use the equations and to make predictions, but there is no agreement yet on how nature produces the effects described by quantum mechanics. Since we have no satisfactory answer to this question, we have to collect the available arguments for a future consistent picture which explains the appearance of quantum phenomena.

Most of the community is satisfied with wave-particle duality, based on the experience that quantum phenomena can be explained by waves, but detection at low intensities is registered locally as being produced by particles. The question is raised whether this is related to the atomic structure of detectors. On the other hand, some authors are inclined to the explanation by wave phenomena. They derive the Schrödinger equation from the wave equation as an approximation for potentials much smaller than the rest mass, see e.g., Section 6.5 of Ref. [4]. They explain the Born rule by the quadratic Hamiltonian for electromagnetism and the linear superposition principle from the linear equations, which are valid for waves.

Waves and particle approaches do not take into account that there are phenomena between waves and particles, as well as topological solitons, which are formulated as field models. By their topological properties, special waves characterised by topological quantum numbers become spatially concentrated. These properties of topological solitons may explain why we observe both wave and particle properties. An interesting experiment which demonstrates how wave and particle properties may cooperate is the oil drop experiment invented by Couder [6]. A numerical simulation of travelling waves, solitons, and radiating excess energy as dispersive waves is described in Sect. VI.5 of the book by Komech [7].

In this article, I intend to offer another insight into quantum mechanical processes by employing the maximal entropy random walk (MERW), which is one of the approaches

in studying diffusion processes in complex systems [8–11]. It models processes where the stationary probability of finding a particle localises in the largest nearly spherical region that is free of defects. Thus, MERW differs on irregular lattices from the generic random walk (GRW), where, at each time step, particles go with equal probabilities to any one of its nearest neighbours. This allows us to derive the stationary Schrödinger equation, to explain the Born rule [12–14] and, as shown in this article for the first time, to obtain the Darwin term, which usually appears in a relativistic formulation only [15,16]. The rules of linear algebra follow as a consequence of the solution of a variational problem.

MERW is a Markov chain, its time evolution of trajectories depends on the present state only and not on the past. As discussed in Section 2, it is defined by a step matrix *M*, acting on vectors of position distributions, known as amplitudes. In distinction from other random walks, it is also defined by the additional request that all paths taking the same time are assumed to have the same probability. Such a uniform distribution of trajectories has a maximal entropy and justifies the choice of the name [8]. For particles in a potential, the solutions of the variational problem are given by the eigenfunctions of *M*. They provide the ensemble of paths and its equilibrium distribution. Measurements are conducted between the equilibrium distributions in the past and future. From the identity of these distributions follows the time symmetry of the ensemble of trajectories forward and backward in time and the Born rule. In Section 3, we apply MERW to the free motion in a box and to motion in a potential. As long as the free motion is not inhibited by boundaries or a potential and if the diffusion constant is determined under the assumption that in reduced Compton, time particles may move a distance of a reduced Compton wavelength, we find that the amplitudes obey the time dependent Schrödinger equation continued in imaginary time. In the case of the influence of a potential, we find that the infinite time limit, the stationary distribution, corresponds to the solution of the time-independent Schrödinger equation. Expanding the solution of the variational problem up to α_f^4, we obtain the Darwin term. In Section 4, we discuss the problems with the probability interpretation of excited states and their technical solution.

We formulate MERW on a cubic homogeneous lattice with spacing δ. Expanding to the second order in δ, we can compare the terms with the three terms which we obtain in the nonrelativistic expansion of the Dirac equation. Despite the nonrelativistic formulation, MERW can describe Zitterbewegung and obtain the correct form of the Darwin term. From the nonrelativistic treatment, it is understandable that the other two terms differ.

2. Method

In Refs. [8,12], MERW was introduced as a stochastic process in discrete space and time on general graphs, which are structures made of vertices (nodes) x and edges xy. In a single time step of a Markov chain, the particle may hop along an edge from x to a neighbouring node y with a probability described by the stochastic (transition; Markov) matrix S_{xy} independent of the past history, leaving the probability distribution ρ_x invariant; see Equation (9). Equal probabilities S_{xy} outgoing from a vortex x to all neighbours y define the commonly used generic random walk (GRW). Due to this equality, GRW maximises the entropy locally, whereas we Will define MERW such that it maximises the entropy of the ensemble of paths of an equal length.

Starting with an initial distribution $|i\rangle$ of positions MERW counts with the amplitudes

$$\psi_i(x) := \langle x|i\rangle \qquad (1)$$

the number of paths starting at x. MERW is based on a step matrix M with non-negative matrix elements $\hat{M}(y,x) := \langle y|M|x\rangle$ which is defined by

$$|f\rangle = M^k|i\rangle \quad \Leftrightarrow \quad \psi_f(y) = \sum_x \hat{M}^k(y,x)\psi_i(x), \qquad (2)$$

the number of paths ending after k time steps at $|y\rangle$. The elements of M are defined in Equation (15) for particles in a box and in Equation (33) for particles in a potential V.

Constructing the possible paths with this step matrix M guarantees that all trajectories which take the same time have an equal probability. This uniform distribution of paths has a maximal entropy and justifies calling the method "Maximal Entropy Random Walk". The principle of maximum entropy [17] states that the probability distribution which best represents the current state of knowledge about a system is the one with the largest entropy. For time-reversal invariant processes, we can assume that steps away and back are equally likely and thus that the matrix M is symmetric.

Our interest concentrates on the stationary states of this diffusion process. We obtain them from the eigenvalue equation of the symmetric step matrix M

$$M|i\rangle = \lambda_i |i\rangle \quad \Leftrightarrow \quad \underbrace{\langle x|M|i\rangle}_{\hat{M}\psi_i(x)} = \lambda_i \underbrace{\langle x|i\rangle}_{\psi_i(x)}. \tag{3}$$

For any finite number of time steps, these states are normalisable and define orthogonal rays in a real vector space

The set of orthonormal eigenvectors $|i\rangle$

$$\langle i|j\rangle = \delta_{ij} \tag{4}$$

defines a complete basis

$$\sum_{i=1}^{N} |i\rangle\langle i| = \mathbb{1}_n. \tag{5}$$

If the limit for an infinite number of steps corresponds to a bound state, this limit is also normalisable.

Due to the request of the highest entropy of the ensemble of paths, the state $|0\rangle$ with the highest eigenvalue λ_0 is stable,

$$\langle x|0\rangle\langle 0|y\rangle = \lim_{k\to\infty} \sum_i \langle x| \underbrace{\left(\frac{\lambda_i}{\lambda_0}\right)^k}_{\to \delta_{i0}} |i\rangle\langle i|y\rangle \overset{(3)}{=} \lim_{k\to\infty} \langle x| \left(\frac{M}{\lambda_0}\right)^k \sum_i |i\rangle\langle i|y\rangle =$$

$$\overset{(5)}{=} \lim_{k\to\infty} \langle x| \left(\frac{M}{\lambda_0}\right)^k |y\rangle, \tag{6}$$

the other eigenstates are metastable. The Perron–Frobenius theorem in the matrix theory states for which matrices there exists a highest real eigenvalue λ_0 with non-negative components $\langle x|0\rangle \geq 0$, i.e., without nodes.

Due to the iterability

$$\langle x|0\rangle \overset{(3)}{=} \langle x| \frac{M}{\lambda_0} \underbrace{\sum_y |y\rangle\langle y|}_{1} 0\rangle \overset{(3)}{=} \langle x| \frac{M}{\lambda_0} \underbrace{\sum_y |y\rangle\langle y|}_{1} \frac{M}{\lambda_0} \underbrace{\sum_z |z\rangle\langle z|}_{1} 0\rangle \tag{7}$$

we are able to find the stochastic matrix S_{xy}. Stochastic matrices are also called transition or Markov matrices and determine the transition probabilities of Markov chains. As the matrix elements are probabilities, their values satisfy

$$0 \leq S_{xy} \leq 1. \tag{8}$$

According to Equation (10), the matrices S are leftstochastic, this means that their columns form probability vectors and thus add up to 1; see Equation (10). The 1-norm is $||S||_1 := \max \sum_y |S_{xy}| = 1$. Since the spectral radius of a matrix is smaller than the norm,

all eigenvalues of S are smaller or equal to 1, which leaves a probability distribution ρ_x invariant

$$\rho_z = \sum_x \rho_x S_{xz} = \sum_x \rho_x \sum_y S_{xy} S_{yz} \tag{9}$$

and with a sum over final states of value 1 from a given initial site x

$$\sum_y S_{xy} = 1. \tag{10}$$

Equations (9) and (10) can be fulfilled by

$$\rho_x := \frac{\langle 0|x\rangle\langle x|0\rangle}{\langle 0|0\rangle}, \quad S_{xy} := \frac{1}{\langle x|0\rangle}\langle x|\frac{M}{\lambda_0}|y\rangle\langle y|0\rangle, \tag{11}$$

where the denominators normalise the number of initial positions in the ground state to 1. We obtain, for the iterated stochastic matrix

$$\sum_y (S^k)_{xy} \overset{(11)}{=} 1 \tag{12}$$

and the condition

$$\sum_x \rho_x \overset{(11)}{=} \frac{\langle 0|\overbrace{\sum_x |x\rangle\langle x|}^{\overset{(5)}{=}1}|0\rangle}{\langle 0|0\rangle} \overset{(5)}{=} 1 \tag{13}$$

which is requested for a probability distribution.

The entropy growth in MERW processes is explicitly defined in Appendix A. There are also examples to show that the entropy is maximal for the states with the highest eigenvalue λ_0.

An interesting discussion about the interpretation of the wave function in quantum mechanics, as the probability amplitude that the system is in a certain configuration, can be found in the works of Mara Beller [18].

3. Applications of MERW
3.1. Free Motion in a Box

The first impression about the method that we have for free motion is in a box of size L^D and infinitely high walls in D dimensions. We are using the sites $\vec{x}$, coordinates x_d, unit vectors $\hat{d}$ and lattice spacing δ for $(N+1)^D$ sites

$$x_n^d = \left(n^d - \frac{N}{2}\right)\delta, \quad n^d \in \{0, 1, 2 \ldots N\}, \quad \delta = \frac{L}{N}. \tag{14}$$

We describe free motion with a step matrix M, e.g., by

$$\langle\vec{x}|M|\vec{y}\rangle = \langle\vec{y}|M|\vec{x}\rangle = \begin{cases} 1 & \text{for} \quad |\vec{x}-\vec{y}| = \delta, \quad \vec{x} \text{ and } \vec{y} \text{ not at the boundary,} \\ 0 & \text{else.} \end{cases} \tag{15}$$

We can start the time evolution with some small initial distribution around $\vec{\mu}$, somewhere in the middle of the lattice, e.g.,

$$\mathring{\psi}(\vec{x}) := \prod_{d=1}^{D} \delta_{x_d,\mu_d}, \tag{16}$$

and evolve the amplitudes by

$$\psi(\vec{x}, \tau) := \sum_{\vec{y}} \hat{M}^\tau(\vec{x}, \vec{y}) \, \mathring{\psi}(\vec{y}). \tag{17}$$

According to the central limit theorem, the amplitudes tend asymptotically to the Gaussian distributions

$$\phi(\vec{x}, \tau) \propto \exp\left\{-\frac{(\vec{x} - \vec{\mu})^2}{2\sigma_\tau^2}\right\}, \tag{18}$$

as long as the distribution does not reach the boundaries.

Due to the additivity of variances, the variance of a coordinate increases in every time step by the same amount $\delta\sigma^2$ leading after τ time steps to the variance

$$\sigma_\tau^2 \overset{(20)}{=} \tau \, \delta\sigma^2. \tag{19}$$

For the step matrix M of Equation (15), we obtain

$$\delta\sigma^2 \overset{(15)}{=} \frac{2\delta^2}{2D} = \frac{\delta^2}{D}. \tag{20}$$

Normal diffusion, obeying Fick's law, would approach a constant distribution, whereas MERW ends in the state of the highest entropy, with an amplitude proportional to a product of cosine functions

$$\psi_0(x) \propto \prod_{d=1}^{D} \cos(k_0 x^d) \quad \text{with} \quad k_0 = \frac{\pi}{L}. \tag{21}$$

To check this behaviour, we have evolved, in Figure 1, amplitudes in one dimension for $L = 1$, $N = 32$ and $\mu = 0$, according to Equation (17). In the left diagram, we compare the amplitudes $\psi(x, 16)/\psi(0, 16)$ (blue dots), after 16 time steps, before the distribution arrives at the boundary, with the corresponding Gaussian $\phi_{16}(x)/\phi_{16}(0)$ of free diffusion (red line). We find a similar nice agreement in the right diagram after 256 time steps between $\psi(x, 256)/\psi(0, 256)$, with red dots, and the quantum mechanical ground state $\cos(\pi x)$, the blue line, in a box of size 1. According to Equation (15), there are no paths leading into the walls. This is why we observe in the diagram on the right that particles cannot accumulate near the walls where the freedom of movement is restricted. There are many more paths that lead away from the walls, into areas where particles can move freely.
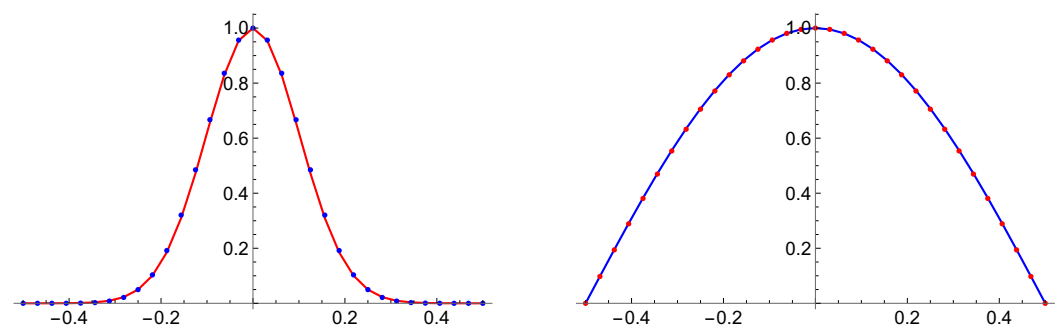

Figure 1. (**Left**): The blue dots depict the amplitudes $\psi(x, 16)/\psi(0, 16)$ according to Equation (17) and the red line joins the dots of $\phi_{16}(x)/\phi_{16}(0)$ according to Equation (18). (**Right**): After 256 time steps $\psi(x, 256)/\psi(0, 256)$ (red dots), the cos-function approaches $\cos(\pi x)$, as shown by a blue line.

The normalised time-dependent Gaussian

$$\phi(\vec{x}, \tau) := \frac{1}{(2\pi \, \tau \, \delta\sigma^2)^D} \exp\left\{-\frac{\vec{x}^2}{2\tau \, \delta\sigma^2}\right\} \tag{22}$$

is a solution of the diffusion equation

$$\partial_\tau \phi(x, \tau) = D_d \, \Delta \, \phi(x, \tau) \quad \text{with} \quad D_d = \frac{\delta \sigma^2}{2} \tag{23}$$

with the diffusion constant D_d. Introducing the imaginary time $\tau = \mathrm{i} \frac{t}{\delta t}$ and the time step δt

$$\mathrm{i} \partial_t \phi(x, t) = -\frac{D_d}{\delta t} \, \Delta \, \phi(x, t). \tag{24}$$

the diffusion equation agrees with the time-dependent Schrödinger equation for a free particle [19], if we choose

$$D_d := \frac{\hbar}{2m} \, \delta t \tag{25}$$

leading to

$$\frac{\hbar}{m} \overset{(25)}{\underset{(23)}{=}} \frac{\delta \sigma^2}{\delta t}. \tag{26}$$

This fraction can be expressed by the ratio of the square of the reduced Compton wave length $\bar{\lambda}_c$ and the reduced Compton time $\bar{t}_c$

$$\frac{\hbar}{m} = \frac{\bar{\lambda}_c^2}{\bar{t}_c} \quad \text{with} \quad \bar{\lambda}_c := \frac{\hbar}{mc}, \ \bar{t}_c := \frac{\hbar}{mc^2} \tag{27}$$

Inserting Equation (26), we conclude

$$\frac{\delta t}{\bar{t}_c} \overset{(26)}{\underset{(27)}{=}} \frac{\delta \sigma^2}{\bar{\lambda}_c^2}. \tag{28}$$

In the n^2 reduced Compton times, the diffusion therefore propagates in every direction n and reduced Compton wave lengths; the single step is conducted with velocity c_0/n. What is appropriate for quantum physics could be $n = 1/\alpha_f = 137.036$, where α_f is Sommerfeld's fine structure constant. Since the further results are not influenced by the factor n, we choose further results for simplicity $n = 1$, and obtain

$$\delta t = \bar{t}_c = \frac{\hbar}{mc^2}, \quad \delta \sigma^2 = \bar{\lambda}_c^2 \overset{(27)}{=} \frac{\hbar^2}{m^2 c^2}. \tag{29}$$

After reaching the stationary solution, depicted in the right diagram of Figure 1, the time derivative in the diffusion Equation (23) vanishes, and the equation agrees with the time-independent Schrödinger equation. The solution is now influenced by boundary conditions. The problem has turned to a boundary value problem, characterised by eigenvalues and eigenfunctions, and by rays in a vector space.

3.2. Motion in a Potential

From free motion, we had the idea that the solutions of MERW may agree with solutions of the stationary Schrödinger equation. In this section, we will investigate under which conditions we can obtain from MERW the stationary Schrödinger equation for a particle in a potential.

In classical and quantum mechanics, we use time as a universal parameter, despite knowing that time in different gravitational potentials runs with a different speed [20]. With the concept of energy as a factor multiplying this universal time [21], we describe moving particles measuring the action $S = \int \mathrm{d}t L$ in units of $\hbar$. In a potential V and for vanishing kinetic energy T, the Lagrangian reads

$$L := T - V = -V \quad \text{for} \quad T = 0. \tag{30}$$

Since in MERW all paths that take the same amount of time are weighted equally, we have to take into account in MERW that the parameter time changes at different rates in different environments. The time evolution for a time step δt is described by the formal Taylor series

$$\psi(x, \tau + \delta t) = \exp\{\delta t\, \partial_\tau\}\psi(x, \tau). \tag{31}$$

According to classical mechanics, the action S generates the canonical transformations between initial coordinates and momenta and their time-dependent values. Using this relation, for the step matrix, we obtain

$$\hat{M} \overset{(31)}{\underset{(2)}{=}} \exp\{\delta t\, \partial_\tau\} = \exp\{\frac{S(\tau, \tau + \delta t)}{\hbar}\} \overset{(30)}{=} \exp\{-\frac{V \delta t}{\hbar}\} \overset{(29)}{=} \exp\{-\frac{V}{mc^2}\}. \tag{32}$$

We work in an infinite cubic lattice in R^D with lattice spacing δ. The variation of the potential along the path in direction $\hat{d}$ is considered by the mid-point rule, well known from the path of integral formulation of quantum mechanics [22].

$$\hat{M}_d(\vec{x}) := \exp\{-\frac{V(\vec{x} + \hat{d}\delta/2)}{mc^2}\} \in \mathbb{R}_+, \quad d \in \{1, 2, \ldots, D\}. \tag{33}$$

In every time step, determined by the Compton frequency, the particle has to move to a nearest neighbour. Under this assumption, the eigenvalue Equation (3) for the D-dimensional problem reads

$$e^{\frac{V(\vec{x})}{mc^2}} \lambda_i \psi_i(\vec{x}) \overset{(3)}{=} e^{\frac{V(\vec{x})}{mc^2}} \langle \vec{x}|M|i\rangle =$$

$$\overset{(33)}{=} \sum_{d=1}^{D}\left[\psi_i(\vec{x} + \hat{d}\delta)\, e^{\frac{V(\vec{x}) - V(\vec{x} + \hat{d}\delta/2)}{mc^2}} + \psi_i(\vec{x} - \hat{d}\delta)\, e^{\frac{V(\vec{x}) - V(\vec{x} - \hat{d}\delta/2)}{mc^2}}\right]. \tag{34}$$

To obtain potential differences on the right side of this equation, we multiplied Equation (3) by the factor $e^{\frac{V(\vec{x})}{mc^2}}$.

We expand Equation (34) to increasing orders of δ and obtain in the zeroth order an equation for the approximate value of the eigenvalue

$$e^{\frac{V(\vec{x})}{mc^2}} \lambda_i = 2D + \mathcal{O}(\delta^2) \tag{35}$$

which is not really solvable before knowing the effective potential for the eigenvalue λ_i. We will obtain this value at order δ^2 only. We expand the expressions on the right side of Equation (34) up to order δ^2

$$\psi_i(\vec{x} \pm \hat{d}\delta)\, e^{\frac{V(\vec{x}) - V(\vec{x} \pm \hat{d}\delta/2)}{mc^2}} =$$

$$= [\psi_i(\vec{x}) \pm \delta\partial_d\psi_i + \frac{\delta^2}{2}\partial_d^2\psi_i][1 \mp \frac{\delta}{2}\frac{\partial_d V}{mc^2} - \frac{\delta^2}{8}\frac{\partial_d^2 V}{mc^2} + \frac{\delta^2}{8}\left(\frac{\partial_d V}{mc^2}\right)^2] + \mathcal{O}(\delta^3) \tag{36}$$

and get for Equation (34)

$$e^{\frac{V(\vec{x})}{mc^2}} \lambda_i \psi_i(\vec{x}) \overset{(34)}{=} \sum_{d=1}^{D}\left[\psi_i(\vec{x} + \hat{d}\delta)\, e^{\frac{V(\vec{x}) - V(\vec{x} + \hat{d}\delta/2)}{mc^2}} + \psi_i(\vec{x} - \hat{d}\delta)\, e^{\frac{V(\vec{x}) - V(\vec{x} - \hat{d}\delta/2)}{mc^2}}\right] =$$

$$\overset{(36)}{=} 2D\psi_i + \delta^2\left[\Delta - \frac{\vec{\nabla}V}{mc^2}\vec{\nabla} - \frac{1}{4}\frac{\Delta V}{mc^2} + \frac{1}{4}\left(\frac{\vec{\nabla}V}{mc^2}\right)^2\right]\psi_i(\vec{x}). \tag{37}$$

In this equation, there appear two physical scales, the energy scale mc^2 and the length scale δ. Their relation depends on the physical situation. For hydrogen, we expect the well-known relations of a dimensional analysis

$$E_N = -\frac{mc^2}{2}\frac{\alpha_f^2}{N^2}, \quad r_N = \frac{\hbar}{mc}\frac{N^2}{\alpha_f}, \quad V(r_N) = -\frac{\alpha_f \hbar c}{r_N} = -mc^2\frac{\alpha_f^2}{N^2} \tag{38}$$

where α_f is Sommerfeld's fine structure constant. We conclude that the dimensionless ratios

$$\frac{V(r_N)}{mc^2} = \mathcal{O}(\alpha_f^2) \quad \text{and} \quad \frac{\delta^2}{r_N^2} \overset{(20)}{\underset{(29)}{=}} \frac{D(\frac{\hbar}{mc})^2}{r_N^2} = \mathcal{O}(\alpha_f^2) \tag{39}$$

are both of the order α_f^2 allowing us to expand Equation (37) to an increasing order in α_f^2.

We reach order α_f^2

$$\lambda_i\,\psi_i(\vec{x}) \overset{(37)}{\underset{(29)}{=}} (1 - \frac{V}{mc^2})\,2D\,\psi_i + D\frac{\hbar^2}{m^2 c^2}\,\Delta\psi_i. \tag{40}$$

After multiplication with $mc^2/(2D)$, we obtain the stationary Schrödinger equation for a particle in a potential $V(\vec{x})$

$$\left\{-\frac{\hbar^2}{2m}\Delta + V(\vec{x})\right\}\psi_i(\vec{x}) \overset{(40)}{=} \underbrace{mc^2\left(1 - \frac{\lambda_i}{2D}\right)}_{E_i}\psi_i(\vec{x}) \tag{41}$$

and can read off the relation between energy eigenvalues E_i and eigenvalues λ_i of the step matrix M. The higher the entropy, the lower the energy.

Expanding Equation (37) up to α_f^4, i.e., the exponential function up to the second order in $\frac{V}{mc^2}$, we obtain corrections to the Schrödinger equation

$$\left\{-\frac{\hbar^2}{2m}\Delta + V(\vec{x}) \overbrace{-\frac{V^2(\vec{x})}{2mc^2} + \frac{\hbar^2}{2m^2c^2}V(\vec{x})\Delta + \frac{\hbar^2}{2m^2c^2}\vec{\nabla}V\vec{\nabla} + \frac{\hbar^2}{8m^2c^2}\Delta V(\vec{x})}^{H_{\text{corr}}}\right\}\psi_i(\vec{x}) = \overset{(37)}{\underset{(29)}{=}} \underbrace{mc^2\left(1 - \frac{\lambda_i}{2D}\right)}_{E_i}\psi_i(\vec{x}). \tag{42}$$

It is interesting to compare the correction terms H_{corr} with the correction terms derived from the Dirac equation [23]

$$H_{ck} = -\frac{\mathbf{p}^4}{8m^3c^2} = -\frac{T^2}{2mc^2}, \quad H_{\ell s} = \frac{1}{2m^2c^2}\frac{1}{r}\frac{dV}{dr}\boldsymbol{\ell}\mathbf{s}, \quad H_D = \frac{\hbar^2}{8m^2c^2}\Delta V. \tag{43}$$

The first term H_{ck} is the relativistic correction to the kinetic energy, the second term $H_{\ell s}$ is the spin-orbit term. The third term H_D is the Darwin term and usually attributed to Zitterbewegung of the electron.

We do not expect a relativistic correction to the kinetic energy in MERW, as it is a nonrelativistic model. Indeed, the corresponding first two terms in H_{corr} of Equation (42) cancel. This can be observed, using the virial theorem

$$V = -2T, \tag{44}$$

which is valid for the hydrogen atom already in the classical description. Further, using the quantum mechanical operator for the kinetic energy

$$T = -\frac{\hbar^2}{2m}\Delta \tag{45}$$

it follows

$$\frac{V^2(\vec{x})}{2mc^2} \stackrel{(44)}{=} \frac{2T^2(\vec{x})}{mc^2} \stackrel{(45)}{\underset{(44)}{=}} \frac{\hbar^2}{2m^2c^2}V(\vec{x})\Delta. \tag{46}$$

This shows that the relativistic correction to the kinetic energy vanishes in MERW, as it should. The spin-orbit term has some similarity to the third term in H_{corr}, as both contain the gradient of the potential. But, as expected, they cannot really agree since MERW does not include spin degrees of freedom.

The last term of H_{corr} in Equation (42) has exactly the form of the Darwin term

$$\frac{\hbar^2}{8m^2c^2}\Delta V(\vec{x}) = H_D. \tag{47}$$

This term was first derived by Charles Galton Darwin when he solved the Dirac equation for the motion of an electron in a central force field [15]. It can be derived in a general form in a nonrelativistic expansion of the Dirac equation [16]. Usually, it is assumed that this term has no classical explanation [24]; it is interpreted as an electron–positron pair creation with the subsequent annihilation of the electron of hydrogen with the positron—see Figure 2. In MERW, it does not need annihilation processes because it is a consequence of the diffusive motion. About the origin of this diffusive motion, one can only speculate.
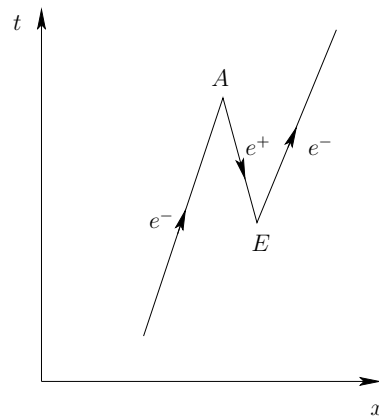

Figure 2. Schematic picture of Zitterbewegung of electrons. An electron e^- propagates in x-direction. Then, after some time, in E, an electron–positron pair is created. Electrons e^- are represented by arrows in time direction and the positron e^+ by arrows backward in time t. At A, the original electron annihilates with the positron. By this process, the position of the original electron appears to be shifted by some accidental distance to a new position.

4. Excited States

As mentioned above, amplitudes of states with the highest entropy, quantum mechanical ground states, are non-negative and allow for an interpretation as numbers of trajectories. The case of a particle in a box with infinite walls shows that the increasing energy corresponds to increasing values of the negative Laplacian, see Equation (41), and to an increasing number of nodes, where the amplitudes change their sign. Since amplitudes count numbers of trajectories, this poses a problem in the probability interpretation of trajectories in MERW. Analogous problems can already be identified in the first paper by Schrödinger on wave mechanics [25] and in experiments on bouncing neutrons [26]. Schrödinger, after around two years of intensive deliberation (his notebooks from this time are available), decided to start from the Hamilton–Jacobi equation. He substituted the

action S by the exponential of S. Comparing the experiment, he realised that the natural unit of S is the reduced Planck constant $\hbar$. Therefore, he equated $\psi = e^{S/\hbar}$. Observe that in this first paper he did not introduce the imaginary i. There is a unique map of S to ψ, but the inverse is not true. The nodes of ψ and negative values of ψ cannot be mapped to S. The existence of nodes also inhibits the particle interpretation of quantum mechanics, as one can nicely observe in the experiment with bouncing neutrons [26]. In this experiment, ultra cold neutrons are observed which are bouncing on a horizontal mirror with heights of roughly 30 µm and a horizontal velocity of 6 m/s. The phase space calculations [27] in this experiment clearly show that in an eigenstate of the corresponding Schrödinger equation, the neutrons cannot cross the nodes of the wave function. The general excuse is that the neutrons are never observed in an eigenstate, but always in a superposition. One has to conclude that true eigenstates are never realised in an experiment. On the other side, it is interesting to observe that in the momentum representation, one can clearly see how the neutrons fall with a vertical velocity, increasing proportionally in time towards the reflecting mirror and bouncing back as expected with a constant deceleration; see Figure 20 of Ref. [27].

Since amplitudes in MERW are normalised numbers of trajectories, they should not be negative. At first sight, negative values for amplitudes seem to contradict Kolmogorov's axioms of probability theory. However, there is an exit to avoid this contradiction. One can assume two species of particles with opposite signs of amplitudes. In the time evolution, both types of trajectories evolve towards the nodes and may cross the nodes. In numerical calculations, applying the step matrix M according to Equation (2), one observes that the amplitudes of opposite signs annihilate each other. This nicely corresponds to trajectories of particles and antiparticles that can annihilate, as we observe them in high-energy physics experiments. With the particle–antiparticle interpretation, one can retain Kolmogorov's axioms and obtain a consistent probability interpretation of excited states as particle and antiparticle trajectories.

Of course it looks rather weird that we describe the spatial trajectories of bouncing neutrons in excited states by the oscillation between particle and antiparticle trajectories, whereas the momentum trajectories allow for a homogeneous interpretation. In a soliton interpretation of particles, this change in the interpretation may be related to the fact that one does not know what type of particles is present, as long as the particles do not react with the surroundings [28].

5. Conclusions

By its construction as a Markov chain with a step matrix M, Maximal Entropy Random Walk (MERW) determines the distribution of particle paths $\psi(\vec{y}, \tau) = \sum_{\vec{x}} \hat{M}^\tau(\vec{y}, \vec{x}) \, \mathring{\psi}(\vec{x})$ starting from an initial distribution $\mathring{\psi}(\vec{x})$ and paths of same duration τ. The final distributions approach an equilibrium distribution, the time derivative term of the diffusion equation vanishes, and there is no difference between evolution forward and backward in time; thus, we arrive at the eigenvalue in Equation (42).

We observed many common properties between MERW and the time-independent Schrödinger equation:

1. MERW treats stationary states of a diffusion process resulting from the eigenvalue Equation (3) of the step matrix M. These states can be represented by rays in a Hilbert space.
2. MERW distinguishes between amplitudes $\langle 0|x \rangle$ and probabilities ρ_x. The amplitudes count the number of trajectories arriving in the equilibrium distribution at a certain position. The probabilities indicate the share of trajectories, which, in the equilibrium distribution, pass at x. They are not modified by the Markov process.
3. MERW respects the Born rule. It derives this rule from the relation (11) between the step matrix M and the stochastic matrix S, and the request that S leaves a probability distribution invariant. MERW impressively shows how the Born rule appears from a measurement between an equilibrium state in the past and the same equilibrium

state in the future. This symmetry of the evolution from past to present, and to the evolution from present to future, was already investigated by Erwin Schrödinger in his article [19] on the reversal of the laws of nature. The translation of this article is available in [29].

4. With the appropriate choice of the diffusion constant (25), MERW derives the Schrödinger equation for a free particle.
5. Generalising the observation that time in a gravitational field is running in different heights with a different speed to arbitrary potentials, MERW respects the rule that all trajectories of same duration are counted with equal probability. In this way, MERW allows us to derive the Schrödinger equation for a particle in a potential.
6. From diffusion, MERW derives both the stationary Schrödinger equation and the Darwin term of the nonrelativistic expansion of the Dirac equation.

If we want to use MERW to describe the properties of nature or to explain the background of the validity of the Schrödinger equation, we realise that MERW leaves several questions open or answers them in the wrong way:

1. MERW is a diffusion process which propagates in n^2 the reduced Compton time $\bar{t}_c = \frac{\hbar}{mc^2}$ a distance of n of reduced Compton wave lengths $\bar{\lambda}_c = \frac{\hbar}{mc}$. The width of the ground state of the atomic hydrogen may indicate $n = 1/\alpha_f$ as a realistic order of magnitude. MERW does not explain the mechanism driving this diffusion process. The pair creation with a necessary energy of around 1 MeV may not be the origin of the Darwin term. After observing that in elastic scattering processes the position of the particles is not conserved, the author is rather inclined to think of some yet unknown scattering processes. In this respect, Ref. [2] argues in favour of a zero-point field.
2. The kinetic term $-\frac{\hbar^2}{2m}\Delta$ in Equation (41) is of a statistical origin only. The velocity seems to correspond to the diffusive velocity, as defined in [2] p. 125. In MERW, there is no "classical" kinetic energy, which is characteristic for the classical motion in Bohr's model of the hydrogen atom. In this sense, MERW cannot describe states with a non-vanishing angular momentum or oscillations in a classical potential. Such excited states appear in MERW as metastable states only, in distinction to the Schrödinger equation, where these states are stable.
3. Amplitudes are numbers of trajectories and should be positive. This is true in the state with the highest entropy, corresponding to the ground state. Amplitudes for states with lower entropy have nodes and therefore regions with negative amplitudes. Trajectories traversing the nodes in opposite directions cancel and thus never cross the nodes. The particle–antiparticle picture has to be introduced to explain these cancellations.
4. MERW can explain the subtraction of trajectory numbers by a particle–antiparticle picture only. Therefore, the explanation of interference in Young's double slit experiment remains very questionable.
5. MERW does not explain the violation of Bell's theorem, as quantum mechanics nicely does.

Due to these reasons, quantum mechanics cannot be simply a result of MERW, but MERW may provide some hints on how to proceed, in order to arrive at a better understanding of quantum mechanics and, like Mermin [30], not be satisfied with "shut up and calculate". With MERW, it is possible to derive the ground-state wave functions of particles in a potential and the exact form of the Darwin term. This could indicate that the ground state of quantum systems is determined by diffusive motion only.

Concerning the history of these ideas, we would like to mention that after a footnote of Eddington [31] concerning the probability in Schrödinger's wave mechanics, which is obtained by two symmetric systems of ψ-waves travelling in opposite directions in time, Schrödinger [19,32] tried to find a probabilistic equation close to his wave equation. A probabilistic formulation of processes with initial and final densities was proposed by Bernstein [33] in 1932. Yasue suggested, in 1981, a time-symmetric variational approach

with an action functional [34]. In the formalism of stochastic processes, Zambrini studied Markovian processes of the Bernstein type in two papers and compared them to quantum mechanics [35–37].

Data Availability Statement: The data presented in this study are available on request from the author.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Entropy for Markov Processes

In information theory, 1 Shannon, the unit of the information content is defined as 1 Bit. If in N events the i-th type appears N_i times, then its probability is

$$p_i := \frac{N_i}{N}, \quad N := \sum_i N_i. \tag{A1}$$

With n bits, we can distinguish between

$$N =: 2^n \tag{A2}$$

events. If a low probability outcome really occurs, this has a high information content. The information content is therefore defined by

$$I(p_i) := \text{lb} \frac{1}{p_i} = -\text{lb} \, p_i \tag{A3}$$

and needs at least $I(p_i)$ bits to store it. Multiplying the information content I_i of different outcomes with the number of their occurrences, we obtain the total number of bits

$$n = \sum_i N_i I_i \overset{(A3)}{=} -\sum_i N_i \text{lb} \, p_i, \tag{A4}$$

necessary to store the information on the outcomes. The average number of bits defines the Shannon entropy H

$$H := \frac{n}{N} \overset{(A4)}{=} -\sum_i p_i \, \text{lb} \, p_i. \tag{A5}$$

The entropy of a process is the amount of information that is necessary to describe this process. A Markov process is defined by a probability distribution ρ_x with the usual properties

$$0 \leq \rho_x \leq 1, \quad \sum_x \rho_x = 1 \tag{A6}$$

and by a transition matrix s_{xy}, which is a matrix of probabilities with

$$0 \leq s_{xy} \leq 1, \quad \sum_y s_{xy} = 1. \tag{A7}$$

A step beginning at x leads with the probability s_{xy} to y. According to Equation (A5), at least $-\text{lb} \, s_{xy}$ bits are necessary to store the information about a step from x to y. $k\rho_x$ of k steps start at x and $k\rho_x s_{xy}$ lead from x to y. Therefore, at least $-k \sum_{x,y} \rho_x s_{xy} \, \text{lb} s_{xy}$ bits are

necessary, to store the information about this event. Thus, on average, the information content of a step is

$$H = \sum_x \rho_x \sum_y s_{xy} \, \text{lb} \frac{1}{s_{xy}} \tag{A8}$$

bits. Hence, H is the average information production per step or in brief, the entropy. By definition, the transition probability s_{xy} is the normalised number of paths leading from x to y

$$s_{xy} = \frac{\langle x|M|y\rangle \langle y|\psi\rangle}{\langle x|M|\psi\rangle}. \tag{A9}$$

s_{xy} is undefined if the denominator is zero, i.e., no new path arrives at x. This has no influence on the value (A8) of the entropy, since, for such sites, we have $\rho_x = 0$. For all other sites x, s_{xy} fulfills the requested conditions (A7). Since ψ changes from step to step, s_{xy} is also modified in every step. For an infinite number of steps, $|\psi\rangle$ approaches the stationary solution $|0\rangle$ and s_{xy} converges to S_{xy} of Equation (9). In Figure A1, we evaluate Equation (A8) during the process shown in Figure 1. As is clearly observed, H increases during the process and approaches the entropy for the dominant eigenfunction $|0\rangle$ after an infinite number of steps.
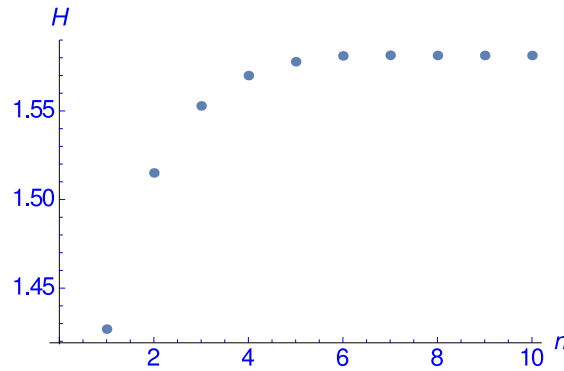

Figure A1. Entropy H for the process discussed in Figure 1 after $k = 2^n$ steps.

For the stationary solution $|0\rangle$, we expect, for k successive steps, an entropy production $H(k) = kH$. This expectation agrees with

$$\begin{aligned}
H(k) &:= \rho_x \sum_{yz} (S^k)_{xy} \, \text{lb} \frac{1}{(S^k)_{xy}} \stackrel{(9)}{=} \rho_x \sum_{yz} S_{xy}(S^{k-1})_{yz} \, \text{lb} \frac{1}{S_{xy}(S^{k-1})_{yz}} = \\
&= -\sum_x \rho_x \sum_y S_{xy} \, \text{lb} S_{xy} \underbrace{\sum_z (S^{k-1})_{yz}}_{\stackrel{(12)}{=} 1} - \sum_y \underbrace{\sum_x \rho_x S_{xy}}_{\stackrel{(9)}{=} \rho_y} \sum_z (S^{k-1})_{yz} \, \text{lb}(S^{k-1})_{yz} = \\
&= -\sum_x \rho_x \sum_y S_{xy} \, \text{lb} S_{xy} - \sum_y \rho_y \sum_z (S^{k-1})_{yz} \, \text{lb}(S^{k-1})_{yz} \stackrel{(A8)}{=} H + H(k-1) = kH.
\end{aligned} \tag{A10}$$

According to Equation (41), the ground state $|0\rangle$ of a particle in a potential V belongs to the highest eigenvalue λ_0. As discussed in Section 2, the eigenvector $|0\rangle$ of M belonging to this highest eigenvalue λ_0 is also characterised by the largest entropy production H per step of Equation (A8). The value of H for a free particle in a linear box $[-0.5, 0.5]$ with 65 sites is depicted in the left diagram of Figure A2, in the neighbourhood of the state $|0\rangle$ for

$$|\psi\rangle = \sqrt{1 - \alpha^2 - \beta^2}|0\rangle + \alpha|1\rangle + \beta|2\rangle, \quad \alpha^2 + \beta^2 \leq 0.06, \tag{A11}$$

where $|1\rangle$ and $|2\rangle$ are the first and the second excited state. As expected, the maximum of H is at $\alpha = \beta = 0$.

With increasing values of α and β, the mixed amplitude (A11) can obtain nodes. This is the reason why, in the left diagram of Figure A2, the values of α and β are restricted to $\alpha^2 + \beta^2 \leq 0.06$ and to values where no nodes appear. This result suggests that we can determine the entropy of excited states from the stochastic matrix which results from blocking the nodes. With this method, we can show that excited states are saddle points of H. Admixing the ground state $|0\rangle$ to the first excited state $|1\rangle$ increases H. All other admixtures decrease H. In the right diagram of Figure A2, the second excited state $|2\rangle$ contributes to $|\psi\rangle$

$$|\psi\rangle = \alpha|0\rangle + \sqrt{1 - \alpha^2 + \beta^2}|1\rangle + \beta|2\rangle, \quad \alpha^2 + \beta^2 \leq 1. \tag{A12}$$

The lowest point at the boundary of the saddle-shaped surface corresponds to an admixture of the ground state and the highest point corresponds to an admixture of the state $|2\rangle$.
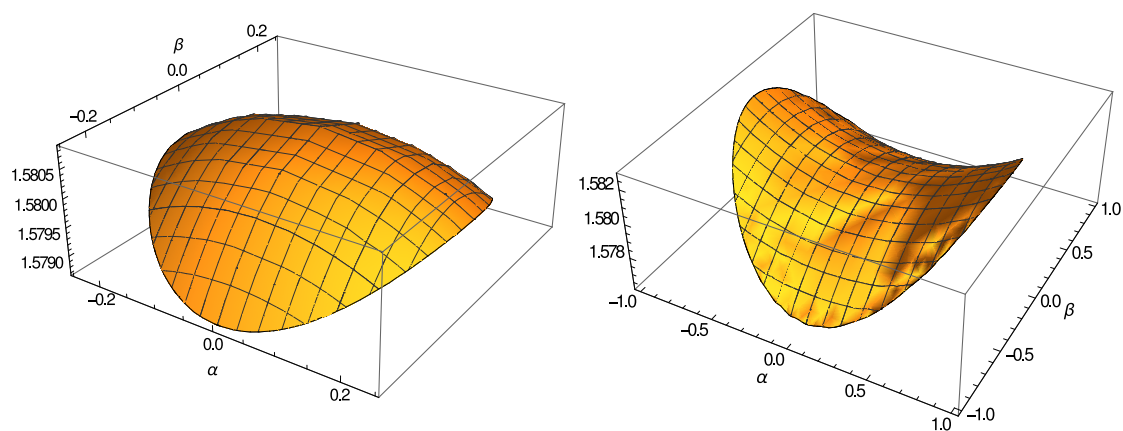

Figure A2. The entropy production H of Equation (A10) per step. In a Markov process, the states move towards higher entropy. Left: For $|\psi\rangle$ of Equation (A11), we observe that the state $|0\rangle$ in the middle of the diagram has maximal entropy. The highest value on the boundary corresponds to an admixture of the state $|1\rangle$ and the lowest point corresponds to that of state $|2\rangle$. Right: For $|\psi\rangle$ of Equation (A12), the state $|1\rangle$ is located at the saddle point in the middle of the diagram. The highest point of the boundary corresponds to the state $|0\rangle$ and the lowest point to the state $|2\rangle$.

As excited states are metastable, a small disturbance is sufficient to let them move towards the ground state, which has the highest entropy.

References

1. Griffiths, D.; Schroeter, D. *Introduction to Quantum Mechanics*; Cambridge University Press: Cambridge, UK, 2018.
2. de la Peña, L.; Cetto, A.; Valdes-Hernandez, A. *The Emerging Quantum: The Physics Behind Quantum Mechanics*; Springer: Cham, Switherland, 2014; p. 366.
3. Cook, D.B. *Schrödinger's Mechanics: Interpretation*; World Scientific (Europe): London, UK, 2018; p. 196. Available online: https://www.worldscientific.com/doi/pdf/10.1142/q0143 (accessed on 6 April 2023). [CrossRef]
4. Ralston, J. *How to Understand Quantum Mechanics*; IOP Concise Physics; Morgan & Claypool Publishers: Williston, VT, USA, 2018. [CrossRef]
5. Dong, S.H. *Factorization Method in Quantum Mechanics*; Fundamental Theories of Physics; Springer: Dordrecht, The Netherland, 2007.
6. Couder, Y.; Fort, E.; Gautier, C.H.; Boudaoud, A. From Bouncing to Floating: Noncoalescence of Drops on a Fluid Bath. *Phys. Rev. Lett.* **2005**, *94*, 177801. [CrossRef] [PubMed]
7. Komech, A.I. *Lectures on Quantum Mechanics and Attractors*; World Scientific: Hackensack, NJ, USA; London, UK; Singapore; Beijing, China; Shanghai, China; Hong Kong; Taipei, Taiwan; Chennai, India; Tokyo, Japan, 2022; p. 272. Available online: https://www.worldscientific.com/doi/pdf/10.1142/12602 (accessed on 6 April 2023). [CrossRef]
8. Burda, Z.; Duda, J.; Luck, J.M.; Waclaw, B. Localization of the Maximal Entropy Random Walk. *Phys. Rev. Lett.* **2009**, *102*, 160602. [CrossRef] [PubMed]

9. Sinatra, R.; Gómez-Gardeñes, J.; Lambiotte, R.; Nicosia, V.; Latora, V. Maximal-entropy random walks in complex networks with limited information. *Phys. Rev. E* **2011**, *83*, 030103. [CrossRef] [PubMed]
10. Ochab, J.K. Maximal-entropy random walk unifies centrality measures. *Phys. Rev. E* **2012**, *86*, 066109. [CrossRef] [PubMed]
11. Ochab, J.K.; Burda, Z. Maximal entropy random walk in community detection. *Pthe Eur. Phys. J. Spec. Top.* **2013**, *216*, 81. [CrossRef]
12. Duda, J. Optimal encoding on discrete lattice with translational invariant constrains using statistical algorithms. *arXiv* **2007**, arXiv:0710.3861.
13. Duda, J. Diffusion models for atomic scale electron currents in semiconductor, p-n junction. *arXiv* **2021**, arXiv:2112.12557.
14. Duda, J. Four-dimensional understanding of quantum mechanics and Bell violation. *arXiv* **2009**, arXiv:0910.2724.
15. Darwin, C.G. The Wave Equations of the Electron. *Proc. R. Soc. London. Ser. Contain. Pap. Math. Phys. Character* **1928**, *118*, 654–680.
16. Sakurai, J. *Advanced Quantum Mechanics*; Always learning, Pearson Education, Incorporated: London, UK, 2006.
17. Jaynes, E.T. Information Theory and Statistical Mechanics. *Phys. Rev.* **1957**, *106*, 620–630. [CrossRef]
18. Beller, M. *Quantum Dialogue: The Making of a Revolution*; University of Chicago Press: Chicago, IL, USA, 1999.
19. Schrödinger, E. Über die Umkehrung der Naturgesetze. *Sitz. Ber. Preuss. Akad. Wissen. Berlin Phys. Math.* **1931**, *8 N9*, 144–153.
20. Chou, C.W.; Hume, D.B.; Rosenband, T.; Wineland, D.J. Optical Clocks and Relativity. *Science* **2010**, *329*, 1630–1633. [CrossRef] [PubMed]
21. Faber, M. Energy as a measure for the elapse of time. *arXiv* **2018**, arXiv:1810.13239.
22. Schulman, L. *Techniques and Applications of Path Integration*; Wiley: Hoboken, NJ, USA, 1996.
23. Thaller, B., The Nonrelativistic Limit. In *The Dirac Equation*; Springer: Berlin/Heidelberg, Germany, 1992; pp. 176–192. [CrossRef]
24. Chen, T.W.; Chiou, D.W. Correspondence between classical and Dirac-Pauli spinors in view of the Foldy-Wouthuysen transformation. *Phys. Rev.* **2014**, *89*. [CrossRef]
25. Schrödinger, E. Quantisierung als Eigenwertproblem (Erste Mitteilung). *Ann. Der Phys.* **1926**, *79*, 361–376. [CrossRef]
26. Abele, H.; Jenke, T.; Leeb, H.; Schmiedmayer, J. Ramsey's Method of Separated Oscillating Fields and Its Application to Gravitationally Induced Quantum Phase Shifts. *Phys. Rev. D* **2010**, *81*, 065019. [CrossRef]
27. Suda, M.; Faber, M.; Bosina, J.; Jenke, T.; Käding, C.; Micko, J.; Pitschmann, M.; Abele, H. Spectra of neutron wave functions in Earth's gravitational field. *Z. FüR Naturforschung* **2022**, *77*, 875–898. [CrossRef]
28. Faber, M. From Soft Dirac Monopoles to the Dirac Equation. *Universe* **2022**, *8*, 387. [CrossRef]
29. Muratore-Ginanneschi, P.; Schwieger, K.E. Schrödinger's 1931 paper "On the Reversal of the Laws of Nature" [Über die Umkehrung der Naturgesetze", Sitzungsberichte der preussischen Akademie der Wissenschaften, physikalisch-mathematische Klasse, 8 N9 144–153]. *Eur. Phys. J.* **2021**, *46*, 28. [CrossRef]
30. David Mermin, N. What's Wrong with this Pillow? *Phys. Today* **1989**, *42*, 9–11. [CrossRef]
31. Eddington, A.S. *The Nature of the Physical World, Chapter X, The New Quantum Theory*; Cambridge University Press: Cambridge, UK: 1928; pp. 201–229.
32. Schrödinger, E. Sur la Theórie relativistie de l'electron et l'interpretation de la mechanique quantique. *Ann. Inst. H. Poincaré* **1932**, *2*, 269–310.
33. Bernstein, S. Sur les liaisons entre les grandeurs aléatoires. *Verhandlungen Des Int. Math.* **1932**, *1.Band*, 288–309.
34. Yasue, K. Quantum mechanics and stochastic control theory. *J. Math. Phys.* **1981**, *22*, 1010–1020. [CrossRef]
35. Zambrini, J.C. Stochastic mechanics according to E. Schrödinger. *Phys. Rev. A* **1986**, *33*, 1532–1548. [CrossRef] [PubMed]
36. Zambrini, J.C. Variational processes and stochastic versions of mechanics. *J. Math. Phys.* **1986**, *27*, 2307–2330. [CrossRef]
37. Zambrini, J. Euclidean quantum mechanics. *Phys. Rev.* **1987**, *35*, 3631. [CrossRef]